Towards a quantitative model of the post-synaptic proteome

Molecular BioSystems

Cite this: *Mol. BioSyst.*, 2011, **7**, 2813–2823

www.rsc.org/molecularbiosystems

PAPER

Towards a quantitative model of the post-synaptic proteome†‡

Oksana Sorokina,*[a] Anatoly Sorokin[b] and J. Douglas Armstrong[a]

Received 20th April 2011, Accepted 17th August 2011
DOI: 10.1039/c1mb05152k

The postsynaptic compartment of the excitatory glutamatergic synapse contains hundreds of distinct polypeptides with a wide range of functions (signalling, trafficking, cell-adhesion, *etc.*). Structural dynamics in the post-synaptic density (PSD) are believed to underpin cognitive processes. Although functionally and morphologically diverse, PSD proteins are generally enriched with specific domains, which precisely define the mode of clustering essential for signal processing. We applied a stochastic calculus of domain binding provided by a rule-based modelling approach to formalise the highly combinatorial signalling pathway in the PSD and perform the numerical analysis of the relative distribution of protein complexes and their sizes. We specified the combinatorics of protein interactions in the PSD by rules, taking into account protein domain structure, specific domain affinity and relative protein availability. With this model we interrogated the critical conditions for the protein aggregation into large complexes and distribution of both size and composition. The presented approach extends existing qualitative protein-protein interaction maps by considering the quantitative information for stoichiometry and binding properties for the elements of the network. This results in a more realistic view of the postsynaptic proteome at the molecular level.

Introduction

Synaptic transmission depends on a very well orchestrated sequence of biochemical processes on both sides of the neuronal synapse. The aggregation of protein complexes of different sizes and composition underpins synapse function, and disruptions at this level are believed to underlie many neuropsychiatric and neurodegenerative diseases.

Proteomic studies suggest the postsynaptic compartment of the excitatory glutamatergic synapse contains up to three thousand distinct polypeptides spanning a wide range of molecular functions.[1–3] The multi-protein signal transduction complex underlying the postsynaptic membrane is referred to as a postsynaptic density (PSD). Its major classes of molecule include receptors, ion channels, cell adhesion proteins and signalling enzymes all brought together and physically linked by diverse scaffold proteins. The resulting, highly stable protein assembly is an electron-dense disc-shaped structure, roughly 40–50 thick, and up to 500 nm wide.[1] It represents a typical example of the so-called 'scaffold-based' signalling complex, where the microenvironment features a highly enriched concentration of signalling components and is condensed into a relatively small sub-cellular volume.[4] Within such complexes, highly conserved, functionally- independent and enriched proteins, assemble transient signalling modules ('signalosomes') through the combinatorial use of common protein interaction domains.[4,5] The properties of any given 'signalosome' will depend on its composition and structural dynamics rather than the activity of any specific component; *i.e.* the complex is more than a simple sum of its parts.

The main components of the PSD are believed to form a lattice-like structure, which provides both basal stability and a mechanism to regulate the signal-dependent structural plasticity of the system. This core lattice structure is based upon precisely defined domain–domain interactions between the several classes of scaffolding proteins. A typical example of such structural domains, highly enriched within the PSD, is the PDZ domain.[6] PDZ domains are often arranged in tandem arrays and often associate with other interaction domains, such as Src homology (SH3) and guanylate kinase (GK) domains to form large multidomain scaffold proteins, such as the membrane associated guanlyate kinases (MAGUKs).[6] The intricate domain composition enables MAGUK proteins to simultaneously bind membrane receptors, channels and cytoplasmic enzymes, bringing together essential elements of

[a] *School of Informatics, University of Edinburgh, Edinburgh, UK. E-mail: oksana.sorokina@ed.ac.uk; Fax: + 44 (0)131 650 6899; Tel: +44 (0)131 6502749*
[b] *Mechanism of Cell Genome Functioning Group, Institute of Cell Biophysics RAS, Pushcino, Moscow region, Russia 142290. E-mail: lptolik@gmail.com; Tel: +7(4967)739319*

† Published as part of a Molecular BioSystems themed issue on Computational Biology: Guest Editor Michael Blinov.

‡ Electronic supplementary information (ESI) available: Supplementary table: Kappa model of PSD; Supplementary Fig. 1. Results of graph analysis for simulated complex. A. Distribution of average path length. B Correlation of average path length with a complex size. See DOI: 10.1039/c1mb05152k.

the signalling cascade.[7] Practically all scaffold proteins in the PSD, including the MAGUKs, GRIP, SHANK and HOMER are able to form homodimers due to self-association mediated by their amino (N)-terminal domains (for MAGUKs), PDZ domains (for GRIP) or SAM domains (for SHANK).[7,8] Since the most of them are also capable of heteromeric association, they can in theory support huge multimeric, multilayer scaffold agglomerations.

One of the most important characteristics of domain–domain protein recognition within the PSD complex is a rather widespread domain ability to bind more than one target sequence motif, or, so-called, domain promiscuity.[9] PDZ domains generally recognize the short conserved peptide motifs located on the C-terminus of the other proteins and are known to cross–react with multiple interaction partners, yet retain some characteristic selectivity within domain subclasses.[9] The same is valid to a greater or lesser extent for other pairs of the complementary domains within the scaffold. For example, SH3 domains can bind proline-rich motifs PXXP, WW domains recognize prolin-rich peptides with consensus PPXY, EVH1 domains associate with proline-rich peptide sequence of type (E/D)FPPPX(D/F).[7,10] It has been proposed that relatively weak binding of proteins *via* PDZ–C-terminus interaction makes the PSD structure very dynamic. In turn, this implies that even weak perturbation can have implications at the level of PSD organisation.[11] At the same time the overall excess of available binding 'slots' guarantees structure integrity and prevents the dissociation of PSD by outward diffusion of components.

In recent years, the availability of high-throughput proteomic and interactomic data has made the analysis of the network representations for the protein complexes a routine task in bioinformatics. Topological analysis of the synaptic interactome reveals the basic principles underlying the functional organization of the protein clusters within the network and helps to identify the most essential elements and network motifs. Several successful studies performed with respect to the PSD demonstrated the modular structure within the complex and linked the protein communities to physiological states of the synapse, giving insight into possible mechanisms of neurological diseases.[2,3]

Although a protein–protein interaction (PPI) graph gives a rather faithful qualitative representation of complex composition, it is challenging to infer the real structure of the multiprotein complex directly from the properties of the interaction network alone. It is also not possible to fully derive a protein-protein interaction network from a domain–domain interaction graph alone. Both cases provide a map of possibilities, whereas abundance of the complex elements and affinities of their bindings are vital to define the structure and stability of the real complex. Bearing this in mind, one might decide to look for the next generation of modelling approaches that supports predictions at the level of complex stoichiometry from data describing a protein-protein interaction network. This goal could be achieved with the help of a relatively new and fast growing modelling approach called rule-based modelling. Rule-based modelling provides a syntax that can be used to formalize protein interactions. Importantly it provides a mechanism to explicitly describe protein/domain binding sites, affinities, state (post-translational modification) and concentrations. It has been deliberately developed to tackle combinatorial complexity, which inevitably emerges in the situation where each of the plethora of system components bears multiple binding sites, which are subjected to posttranslational modifications and have multivalent binding partners. The enormous number of concurrent modifying and binding events in such systems grows exponentially, and simulations rapidly become computationally intractable. In a rule-based approach, each rule defines only what is essential for a particular interaction and omits all the irrelevant context, so-called 'rule decontextualization'. Accordingly, one rule may account for multiple possible states of a model component (agent) that satisfy the rule implicitly, without increasing the size of the model. This helps avoid the combinatorial explosion in models as they scale in size. Rules can be visualised using graphs/contact maps and easily converted to an executable mathematical model that can be simulated using either deterministic or, more often, stochastic algorithms. In this way, the gradual introduction of rule-based systems may help to solve the limitations of static maps, adding the necessary quantitative information onto the existing protein-protein interaction map.

During the past five years several methodologies for rule-based modelling have been proposed: StochSim, MCell, Smoldyn and ChemCell, Kappa and BioNetGen language (BNGL).[12–16] Although principally similar, each language implements its own specific spectrum of features and could be used alternatively or complementary, depending on the purpose of study. These methods have been validated on receptor signalling models each designed with different rule-based modelling techniques, including Tar-receptor-mediated hemotaxis, *FcεRI* - and TCR (T-cell receptor)-mediated responses in immunoreactivity, GPCR (G-protein coupled receptors)-signalling and many others.[16–19] The advantage of rule-based techniques is that the calculation efficiency does not depend upon a size of the network implied by the set of rules. That makes it possible to simulate the formation of the multisubunit signalling complex together with all complexity of the receptor-mediated phosphorylation cascade.

As a first step into quantitative model development, we focussed on examining the steady states that are reachable by the system rather than on dynamics of transition process. We applied the rule-based approach to study the composition of quasi-steady-state protein complexes in postsynaptic density. We created a rule-based model of PSD with the Kappa language.[15,20] Kappa formalism has previously been used to study 'liquidity' of protein agglomerates at equilibrium.[20] Its implemented simulation algorithm is also insensitive to size of generated complexes.[21] For the initial model we focused on interactions between the proteins comprising the core subset of PSD, including key scaffolding proteins and their closest interaction partners, thus, reproducing the core lattice structure. The dynamics of PSD aggregation was described by rules, taking into account protein domain structure and relative protein abundance, where known. The percentage distribution of PSD proteins according to their functional category was reviewed in.[22] The same study also summarised several reports of stochiometric composition of PSD from in the literature for a core set of PSD components.[22–24] Using this information to set initial concentrations of the model elements, we applied a

stochastic calculus to perform the numerical analysis of the relative distribution of sizes of protein complexes, obtained in steady state. As the exact structure of PSD along with values for binding affinities and complex half-lives are not available for the most reactions considered in the model, we ran the global sensitivity analysis to identify subset of parameters (affinity reaction constants) that appears the most important for formation of realistic structure and composition of PSD.

The resulting model allows interrogation of the critical conditions for the protein aggregation to support large complexes. At the same time, the model can capture the effect of mutations, posttranslational modifications, and alternative splice variants on complex structure and size.

Results

Model building and structure

The work presented here represents a proof of concept, which aims to reproduce the basic lattice composition of the post-synaptic density and demonstrate that the modeling approach can be extended to capture dynamic and quantitative processes at the synapse. The PSD is understood to consist of several families of scaffold proteins cross-linked by a series of key domain–domain interactions. These scaffolds bring together the rest of the elements of PSD, namely the neuro-transmitter receptors, elements of various signaling cascades, cell adhesion and cytoskeletal molecules. In this first attempt we have restricted the model to a reduced set of core elements (54 proteins) manually curated from the literature. The model components are shown in a network diagram (Fig. 1). The description of the main components (agents) considered in the model is summarized below.

Scaffolds

The main scaffold components with their characteristic domains are shown in Fig. 2. Our model includes four genes encoding MAGUKS (from the DLG family of proteins): PSD-95 (DLG4/SAP90), PSD93 (DLG2), SAP102 (DLG3) and SAP97 (DLG1). Each has a similar domain composition (Fig. 2). Among them PSD-95 is known to be one of the most abundant (~300 copies/average PSD) and the most stable (with very little turnover during a 30 min period) proteins in the postsynaptic density.[7]

The MAGUKS each possess 3 PDZ domains. The first two PDZs have similar binding properties and are known to interact with NMDA (*N*-methyl-D-aspartate) receptors and Shaker-type K + channels, which underpins their functional surface clustering and proper subunit composition.[25] The third PDZ domain is used to bind cytoplasmic signalling enzymes such as nitric oxide synthase (nNOS), Ras GTPase-activating protein (SynGAP), Rap GTPase-activating protein (SPAR) and some others. SH3 and GK domains operate in a tandem mode, linking the PSD-95 family proteins to the other scaffolds. Guanylate kinase-associated protein (GKAP) mediates the multilayer organization, simultaneously binding the GK domain of PSD-95 family proteins and PDZ domain of another scaffold, situated deeper, ankyrin repeat-containing protein (SHANK).[26] SH3 and GK domains are also capable of binding each other in both intramolecular and intermolecular way, providing the additional integrity to the multiprotein complex.[27] Different groups of PDZ proteins, glutamate-receptor-interacting protein (GRIP) and protein interacting with C kinase 1 (PICK2) interact with AMPA (α-amino-3-hydroxy-5-methyl-4 isoxazole propionic acid) receptors. GRIP proteins contain up to 7 PDZ domains, amongst which PDZ4 and PDZ5 are able to interact with AMPARs. Others can bind ephrin receptors, Ras guanine exchange factors and many other proteins.[26]

SHANK, another master scaffolding protein, is found deeper in the PSDs than the MAGUKs. It is wedged in between the receptor structure on the cell surface and inner elements of actin cytoskeleton. SHANK self-associates through its Sterile alpha-motif (SAM) domains, thus, contributing to the highly-ordered structure of the PSD.[28] Shank interacts

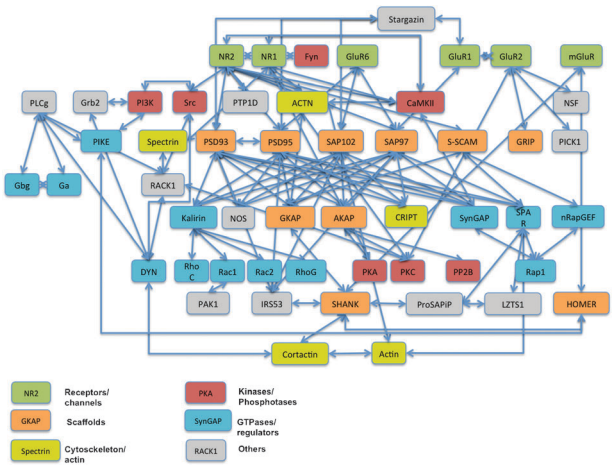

Fig. 1 Protein-protein interaction map of minimal mode of post-synaptic density considered in the Kappa model (54 proteins). Functional protein subcategories are marked by respective colours. This static map captures the interactions in the reduced model of the post-synaptic density.

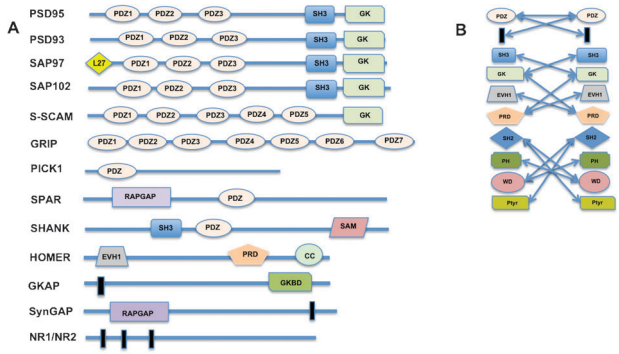

Fig. 2 Domain–domain interaction map considered in the model. A. Domain structure of the major components of PSD. Abbreviations: EVH1, Homer type EVH1 domain; GK, guanylate kinase; L27, L27 domain; PH, pleckstrin homology domain; RapGAP, GTPase-activator protein for Rap/Ran-like GTPases; RasGAP, GTPase-activator protein for Ras-like GTPases; RhoGEF, guanine nucleotide exchange factor for Rho/Rac/Cdc42-like GTPases; SAM, sterile alpha motif; SH3, Src homology 3 domains;.B Complementary domain–domain interacting pairs formalized in the model.

with EVH1 domain of the scaffold protein Homer *via* its prolin-rich motif. The resulting quaternary complex HOMER/SHANK/GKAP/PSD95 is likely to represent a core scaffold structure of PSDs.[8]

Our preliminary model also considers the synaptic scaffolding molecule S-SCAM/MAGI2, which contains 5-6 PDZ domains which can bind the NR2 subunits of NMDArs and guanine-nucleotide exchange factor nRapGEP/RAPGEF2.[29] In addition (like PSD95-related proteins), MAGI2 can bind GKAP *via* its GK domain.[29]

Receptors

NMDARs are the calcium-permeable glutamate receptors of the PSD, playing a central role in synaptic plasticity. They are known to exist as tetrameric complexes composed of two gycine–binding NR1 subunits and two glutamate-binding NR2 subunits.[25,26] NR2 subunits may associate with PDZ1 and PDZ2 domains of PSD-95 family proteins, which favours receptor clustering on the postsynaptic membrane.[25,30]

Another important receptor class, AMPAs, also have a tetrameric structure, mostly presented by GluR1/GluR2 heteromers.[22] GluR2 can bind the PDZ5 domain of scaffold protein GRIP and PDZ domain of PICK1 (protein interacting with C kinase1).[6] GluR1 directly binds SAP97 a member of the MAGUKs family.[31] We included stargazin, the protein which links NMDA and AMPA signalling, binding PDZ domains of PSD-95 related proteins and AMPA receptor subunits.[32] Excitatory synapses are enriched with the other group of receptors: metabotropic glutamate receptors (mGluR1 and mGluR5). In the model we describe co-clustering of mGluRs by means of their simultaneous interaction with EVH1 domain of HOMER and PDZ domain of SHANK.[33]

GTPases and their regulators

Our model incorporates the small GTPase Rap1 and its regulators that are found to bind the PSD9-family proteins through their GK domains. Amongst these is Spine–associated RapGap (SPAR), which stimulates GTPase activity of Rap. Synaptic GTPase-activating protein (SynGAP) has been shown to have dual specificity towards H-Ras and Rap1 and also binds to PSD95-family proteins through one of its three PDZ domains.[34,35] To balance the system we included the guanine-nucleotide exchange factors for Rap1, nRapGEF.[6,36] The MAGUKS also directly bind kalirin-7, a guanine-nucleotide exchange factor for Rho GTPases Rac1, RhoA and RhoG, which is involved in regulation of spine formation *via* actin remodelling.[37–39]

Cytoskeleton

The postsynaptic protein CRIPT, known to influence the microtubule structure, binds the PDZ3 domain of MAGUK scaffolds.[26] The NR1 subunit of NMDA receptors does not interact with PSD95 directly, but instead binds a set of proteins, mainly linked to the cytoskeleton. Among these is α-actinin, an actin cross-linking protein, which links NMDArs to F-actin.[40] Spectrin, which also participates in the actin remodelling, binds to NMDAr independently of α-actinin *via* sites on the NR1 and NR2 subunits.[41] Cortactin, F-actin binding protein, possesses an SH3 domain and is able to interact with prolin-rich domain of SHANK[42] and prolin-rich domain of Dynamins (large GTPases).[43]

Kinases and phosphotases

We have included several of the kinases that dynamically regulate protein interactions within the PSD, and also contribute to multiprotein agglomeration. Among these is Ca^{2+}/calmodulin-dependent protein kinase II (CAMKII), the most abundant protein in PSDs.[22,23] CAMKII exists in a form of a dodecameric oligomer and is believed to play a structural role within a PSD, binding non-competitively with several abundant PSD proteins, such as α-actinin, NR2 NMDA receptor subunit, SynGAP and kalirin.[22,34,40,44,45] MAGUK proteins also interact with A-kinase-anchoring protein 79/150 (AKAP79/150), a scaffold for the serine/threonin kinases PKA and PKC, as well as Ca^{2+}/calmodulin-dependent protein phosphotase (PP2B).[6] Such organization brings the kinases into close proximity with their potential substrates. Among these substrates, NR1 and spectrin are considered in the model as the targets for PKA, and NR1 and GluR1 as the targets for PKC.[41,46] PKC binds receptors for activated C kinase (RACKs), which is also known to perform a scaffolding function, bringing together two distinct kinase types.[47] The multiple WD domains of RACK1 selectively bind the pleckstrin homology (PH) domain containing proteins, such as spectrin and dynamin.[48] At the same time some of these domains are known to be capable of binding Src-family tyrosine kinases, such as Src and Fyn, *via* their SH2 domains. The src family tyrosine kinases are highly expressed in neurons and play a regulatory role in membrane trafficking. In addition to their SH2 domain they also have an SH3 domain, which interacts with proline-rich domains (PRD) of other proteins, such as dynamin and PI3K.[49] Both Src and Fyn can phosphorylate the NR2B subunit of NMDA receptors.[50]

The model also includes the class IA phosphoinositide-3-kinase (PI3K), which is known to participate in synaptogenesis.[51] It interacts with phosphoinositide-3 kinase enhancer (PIKE), which links the PI3K cascade to mGluR receptors, binding the EVH1 domain of scaffold protein Homer *via* a proline-rich domain.[52,53]

We have also included several other proteins that are known to interact with the main scaffold components of PSD like insulin receptor substrate of 53 kDa (IRSp53), nitric oxide synthase (NOS), phospholipase Cγ, ProSAP-interacting protein 1 (ProSAPiP1), LZTS1 and some others.[54–58] The full list of components and associated rules could be found in Supplementary Table 1 and 2.

To date, the model consists of 54 proteins (agents), 136 rules and 84 parameters (Supplementary file 1). All reactions in the system were assumed to take place as concurrent processes within the volume of the spine, which was estimated to be $\sim$ 4.0e–16 L.[22,23] Numerical abundance data for key components, including the PSD-95 family proteins, CaMKII, NMDA and AMPA receptor subunits, SynGAP, GKAP, SHANK, HOMER and some others were sourced from Sheng and Hoogenraad.[22]

Although cooperativity is thought to be essential for operation of some domain tandems (PDZ1/2, SH3/GK, *etc.*),[11] the first model does not take the phenomenon into account, so that all the domains/sites are assumed to work independently. The selected modeling framework can capture this logic and it will

be introduced into future revisions to improve biological plausibility.

The current model assumes mostly unconditional binding, which means that at this point, for simplicity, we generally overlook the post-translational regulation of protein-protein interaction. The modelling framework allows these features and these will be implemented in future studies.

The model considers competitive protein interactions based on domain promiscuity. Accordingly, one of the main model assumptions is that similar domains, which comprise the different peptides, will interact with similar affinity. Even a minimal model of PSD contains 54 components that would give ~150 reversible rules of interaction with ~300 affinity constants. In fact, we assign specific rules for the complementary pairs of domain, such as PDZ1-C-terminal motif, SH3–proline-rich domain (PRD), *etc*. This allows us to reduce (by a factor of 3) the number of constants considered in the model.

Most of the rules in the initial model account for unconditional binding/unbinding and have a form of:

'NR2_DLG_PDZ1' DLG(PDZ1), NR2(c) - > DLG(PDZ1!1), NR2(c!1) @ k8 (1)

'NR2_DLG_PDZ1_diss' DLG(PDZ!1), NR2(c!1) - > DLG(PDZ1), NR2(c) @k_8 (2)

The above two rules in Kappa syntax describe the reactions for association (1) and dissociation (2) for members of MAGUKS/DLG family of proteins family and their interactions with NR2 receptor subunits, where the rate of forwards reaction is k8 and the rate of backward reaction is k_8. In accordance with above, the pair of constants k8/k_8 could be substituted not only for all 4 PSD-95 family members, but also for the rest of the extensive list of model agents that carry PDZ domain and interact with C-terminal motifs of other proteins. While this might seem an over simplification, we considered it appropriate for a demonstration bearing in mind that we can later check for critical parameters and fine tune them. Sensitivity analysis (see below) was performed on the model for this purpose.

The model does not aim to describe the dynamic process of signal propagation through the postsynaptic signaling cascade *per se*, but is rather focused on the steady-state complex association. That is why the regulatory rules, which may condition the binding or unbinding, such as phosphorylation/dephosphorylation of model components were set to an effective minimum.

As a result, dynamic regulation is rather sparsely covered in the current model. However, the elements of the main signaling cascades are included, which gives the perspective for manipulation and demonstrates a proof of principle that we can accommodate such logic. Further revisions to describe most of the regulatory events within consensus PSD, would allow tracking the dynamics of complex aggregation over the time of signal propagation.

Model simulation results

One of the main obstacles to many Systems Biology approaches is the paucity of kinetic data to constrain the model. Ideally we want the exact values for ~100 constants, which are simply not available. However, we found some information for the order and the range of main constants from the literature. For instance, the equilibrium constant (Kd) for the most of PDZ-C-terminus interactions was estimated in a low micromolar range, 1–50 μM.[9] Respectively, the approximate ranges could be obtained from literature for other domain–domain interactions. Some of them would be more specific, others would be estimated by similarity, but in all the cases the value used should not be considered precise. To reduce this uncertainty we optimized the whole system to a biological phenomenon. In our particular case we know that the average PSD has a total

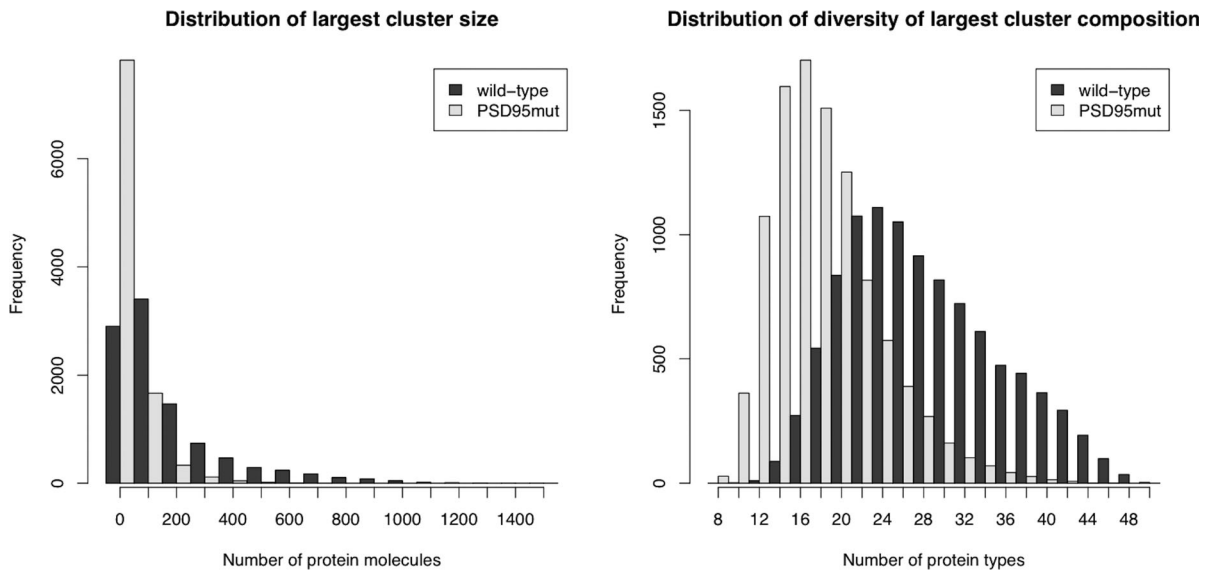

Fig. 3 Results of simulation of Kappa model of the post synaptic density. A. Distribution of sizes (in molecular copies) of the steady- state complexes obtained in 10 000 simulation runs. Different colours correspond to 'wild type' and mutated states of the model (see below) B. Distribution of the diversity of composition of complexes obtained in 10 000 simulation runs. Different colours correspond to 'wild type' and mutated states of the model.

molecular mass of 1.10 ± 0.36 gigadaltons (GDa).[22] Thus, we have reason to assume that we are looking for the set of parameters that gives us a protein agglomeration tending towards that mass. We set the ranges for the each dissociation konstant (Kd) consistent with the available literature data. Also, we set the unimolecular dissociation rates (general unbinding) for all the reactions to the same range of 0.03–0.1 s^{-1}(KappaLanguage.org). After that, 10 000 points (*i.e.* parameter sets) were uniformly sampled using Sobol algorithm from hypercube bound by values for dissociation rate constant and values for dissociation equilibrium constant.

We performed numerical analysis on the relative steady state distribution of protein complexes and their sizes. The results from 10 000 simulations are presented in Fig. 3. Most of the complexes obtained from simulations are relatively small, composed of 200–300 molecules (Fig. 3A). Initial inspection of largest complexes obtained revealed that some of these are simple polymers of scaffold proteins without any channels or receptor attached (not shown). We then examined the molecular composition of complexes. The distribution of protein diversity of the complex components is presented at Fig. 3B. The most common molecular diversity observed was in the range of 25–35 types of molecules per complex. The single biggest complex obtained in simulation, composed of 1586 molecules from 48 types of proteins (Fig. 4B). We also looked at the correlation of size and composition for the complexes. Fig. 3A, shows the index of biggest complex for each of 10 000 simulations *versus* the index of the most diverse complex, and high density on the diagonal provides the evidence that larger complexes in average have more diverse composition.

Supplementary Fig. 2 shows more detailed analysis of simulation results obtained from simulation for one particular parameter set, giving the largest (1586 molecules) complex shown on the Fig. 4B. That particular parameter set gives in total 236 complexes of various sizes from 1 to 1586. Among them, calmodulin participates in more than half of the complexes (114 complexes); PSD95 and SynGap take the second and third place with 76 and 67 complexes, respectively, while GRIP participates in only one complex from 236 obtained. Supplementary Fig. 3 shows the brutto composition of the largest (1586 molecules) complex obtained in this simulation.

Having such a model in hand, it is now simple to simulate situations where stoichiometry of modelled PSD differs from an 'average hippocampal excitatory synapse'. For example, we could easily simulate a PSD 95 knockout. For this, we changed the initial concentration for the PSD95 protein from 300 molecules down to 0. We then ran the simulation in the same manner as it has been described above. The results of this simulation are presented at Fig. 3 and 4. It is clear from the figure that removing PSD95 from the complex distinctly reduced the average sizes of complexes. The most frequently appearing complexes do not exceed the 100 molecules in total (Fig. 3A). The composition of those complexes also became less diverse, with a maximum between 15 to 25 protein types in any simulation (Fig. 3B). The correlation between size and component variability also became less pronounced, especially within the range 0–25 molecule types (X axes) (Fig. 4C). The biggest complex, composed of 729 molecules is presented at Fig. 4D. Interestingly, in this simulation all of the largest complexes comprise of SHANK polymer sheets. Basically, in 'mutant' simulations, the PSD complex seems to recruit additional SHANK molecules replacing PSD95-based polymers in the 'wild-type' simulation. Although we would prefer to extend the model before proposing hypotheses about the nature of the PSD, this is a good example of a readily testable prediction that we can obtain from this kind of modelling strategy. Supplementary Fig. 4 demonstrates proteins participating in the complexes for the example parameter set, giving the largest complex (729 molecules). Here, the RACK1 protein participates in 63 complexes from 212 obtained; calmodullin and actin take the second (54 complexes) and third (41 complexes) places in respect with their popularity. At the same time SHANK is concentrated in single biggest complex as its major component (Supplementary Fig. 5).

We also tested a "double mutant" example, when both concentration of PSD95 and SHANK were set to 0. Simulation showed further reduction both in complex size (most of the complexes are composed of 15–25 molecules) and in complex composition (10–17 protein types). The biggest complex detected in this simulation was composed of 176 molecules. The majority of medium sized (15–25 molecules) complexes consist of other MAGUK-proteins and the first and second order interaction partners (data not shown).

Graph properties of simulated PSD complexes

We next looked at the graph properties of the clusters obtained in simulation. For this we extracted the biggest complexes from all the simulations and analysed them from the point of view of the network topology. We were particularly interested in level of connectivity of the network elements that could be inferred from clustering coefficient, closeness and presence of small bound motifs (cliques).

Using *igraph* package in R[61] we estimated transitivity for each biggest complex, which appeared to be equal to 0 in all cases. As the transitivity reflects the probability of the adjacent vertices of a network to be connected, we can conclude that we are dealing with a rather sparsely connected graph. We also calculated the size of the largest clique (complete subgraph) in each of the biggest networks and found that it never exceeds 2. This supports again our hypothesis that graph is sparse and has a tree-like structure. In comparison, if we calculate the same parameters for the static PPi map, consisting of the model elements, we will get the transitivity equal to 1.15 and maximal cliques of size 4. This means that the highly connected motifs presented at the PPi network are not necessary subjected to inheritance by the rule-generated network and will be transformed with respect to the actual stoichiometry and affinity of network elements.

We also calculated closeness, which represents how many steps are required to access the every other vertex from the given one. The inverse of closeness gives the measure of average shortest path for each vertex. We separately estimated its maximum and minimum values for each cluster as follows:

$$\max cl = \max(1/closeness(g))$$

$$\min cl = \min(1/closeness(g))$$

The *maxcl* value was distributed between 5 and 68. The minimum of *mincl* possessed the value from 2 to 30. Finally,

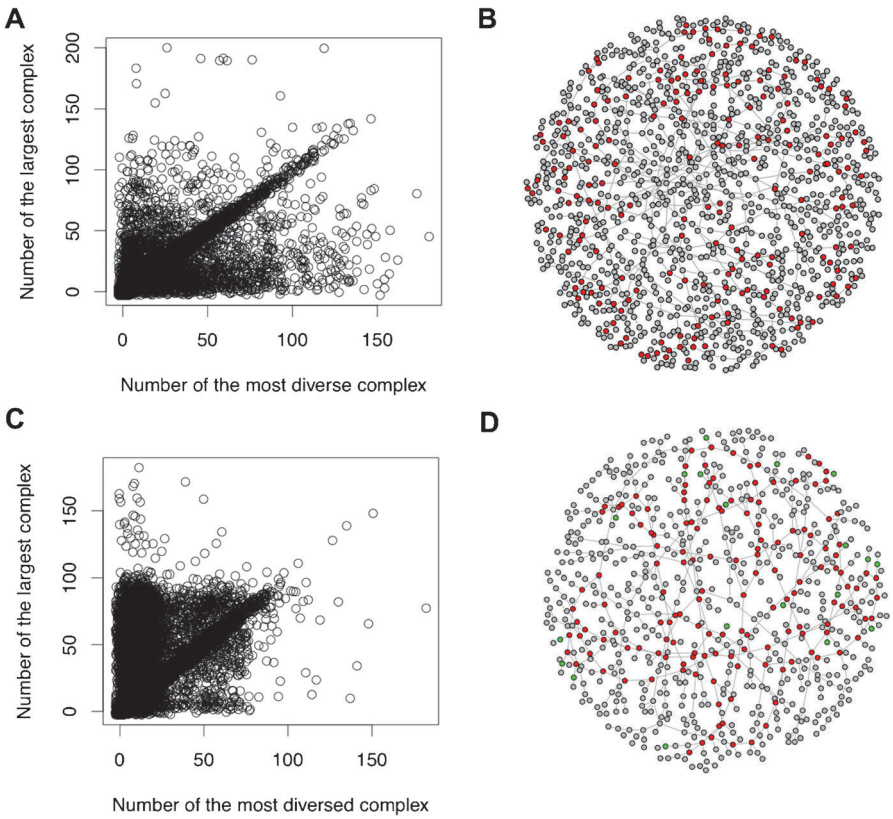

Fig. 4 Results of simulation of Kappa model of the post synaptic density: correlation between size and composition. A. Correlation of size and composition for the complexes obtained in 10 000 simulation runs for 'wild-type' simulation. B. Protein-protein interaction map for the postsynaptic density complex obtained in 'wild-type' simulation (1586 protein copies), comprising the 48 protein types of the 54 considered in the model. Red color corresponds to the MAGUK family proteins. C. Correlation of size and composition for the complexes obtained in 10 000 simulation runs for mutant simulation. D. Protein-protein interaction map for the postsynaptic density complex obtained in PSD95 mutant simulation (729 protein copies). Presented is an example of the largest steady state complex obtained in stochastic simulation. The complex comprises the 32 protein types of the 54 considered in the model. Red color corresponds to the SHANK family proteins.

we estimated the average path, which gives the representation of average coherence of elements of the network. The average path was distributed between 4 and 42 with a mode around 10 (Supplementary Fig. 1).

We looked at the correlation between the size of the complex and its average shortest path and found a few sizable complexes (1300–1500 molecules) with average path about 20 (Supplementary Fig. 1). Among them we chose the complex of 1222 molecules size with the smallest value for *mincl* (12) and marked the correspondent parameter set as potentially giving the most connected structures.

Sensitivity analysis

As mentioned above, we used the same constant values for multiple protein-protein interactions where they are mediated *via* a similar domain–domain association. Although this assumption successfully reduces the parameter set, it might be a critical oversimplification. To address this, we performed sensitivity analysis (PRC coefficient calculated by *sensitivity* package in R[62]), studying the correlation between the values of association and dissociation constants (k*n* and k_*n*) and maximum size of the protein complex obtained in simulation. Fig. 5, A demonstrates the distribution of the constants according to the value of the correlation coefficient. Constants with the largest positive value show positive correlation with the size of the complex, while constants with lowest negative value influence the peptide aggregation negatively. Constants with values around 0 have no measurable effect on the complex size. Constants with the largest magnitude need to be checked more carefully than those tending towards 0, which have minimal effect and could be fixed. We observed k37 having a maximum of positive influence on the complex size. This constant corresponds to coiled-coil domain homodimerization. In the current model it accounts for Homer and ProSAPiP1 protein self-association. Among the most size-affecting constants were observed k23 (dimerization of AMPArs subunits), k33 (α-actinin–F-actin binding), k32 (cortactin–F-actin binding), k27 (PDZ domain of SHANK binding to C-terminus of SHANK interacting proteins), k21 (GK–GK-binding domain interaction), and k29 (SH3–PRD interaction).

We performed the same analysis looking at the complex composition, searching for the constants, having the most influence on the molecular diversity of steady state complexes. The obtained ranking is presented at Fig. 5B. Here, once again, the rates of association and dissociation for coiled-coil domains, k37 and k_37 appear the most important. Other influencing parameters from previous analysis come to the

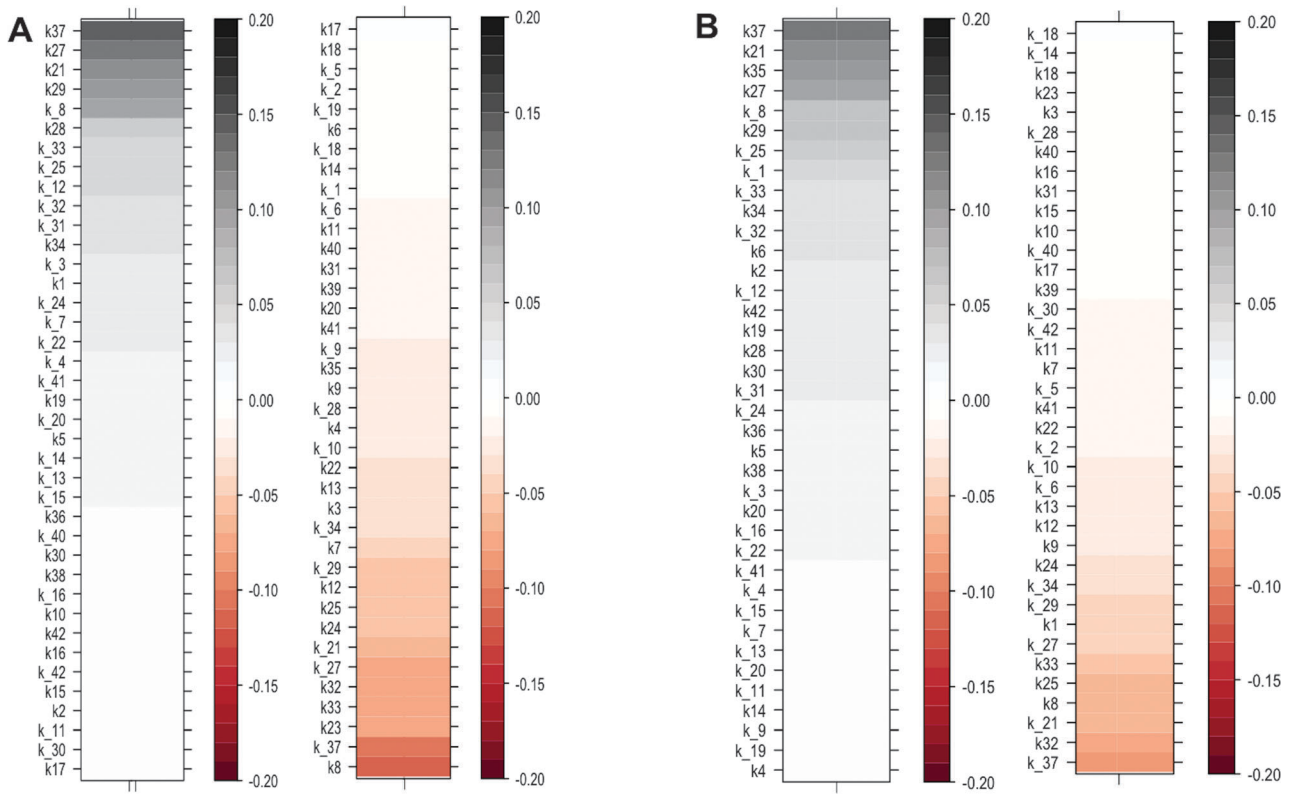

Fig. 5 The results of the sensitivity analysis. A. Constants were ranked according their relative influence on the complex size. The most positively influencing constants are coloured in grey, the most negatively influenced constants are coloured in red. B. Constants were ranked according their relative influence on complex composition. Again, the most positively influencing constants are grey and the negatively influencing-in red.

surface again, though in different order. Additionally, a new parameter, k35 appears in the list. It accounts for the association of stargazin and GluR1, thus forming a physical link between NMDA and AMPA subsystems in the unified complex.

Experimental details

The model consists of 136 rules describing the reversible binding reactions for 54 proteins. Those 54 proteins were selected from the core set of 89 proteins that was detected in most experimental studies of PSD (data not shown). The selection criterion for the 54 proteins was the participation in the scaffold-based lattice structure, which resulted in main scaffold members and their first coordination sphere. The biological information for physical protein association and underlying domain–domain interactions was extracted from the literature (Supplementary table 1). The kinetic information for domain–domain interactions, if available, was also derived from the literature (Supplementary table 1). The proposed rules use the Kappa language semantics (RuleBase.org) and generally have a form:

$$A\ (b),\ B\ (a)\ \langle\text{-}\rangle\ A\ (b!1),\ B\ (a!1),$$

where *A* and *B* are the interacting proteins, a and b are their interacting sites (domains).

Resulting rules have the protein names (PDS95, GKAP, *etc.*) instead of *A* and *B* and domain names instead of *a* and *b*. Domain names used are either traditional (PDZ, SH3, *etc.*) or, if unknown, could be optional, reproducing the name of binding partner.

The model structure was defined in RuleStudio, an Eclipse-based kappa editor.[21,63] The model template was obtained from a a basic interaction map by introduction of symbolic rate constants. The R package randtoolbox[64] was used to sample 10 000 parameter sets from hypercube described. The same R script was also used to create a kappa model corresponding to the individual parameter set obtained from Sobol sequence sampling procedure. Each individual model was simulated by jsim[21] simulator for 1000 s (*i.e.* to steady state). XML files generated by jsim were parsed by an R script with R XML package[65] and final snapshots were converted into igraph[61] representation and then used to create dataset for sensitivity analysis. All source code is available from authors by request.

The adequate sample size (N) should be defined for each model system individually, since it depends on the properties of the system. One way to estimate the optimal N is to systematically increase the sample size and check, whether the set of the most sensitive parameters keeps changing with the increase of N. When two consecutive experiments consistently capture and rank a similar set of most important parameters, one can conclude that there is no evident advantage in further increasing the sample size.

For our network model we used a quantitative metric "top-down coefficient of concordance" (TDCC) to assess the adequacy of the sample size N, as suggested by Marino *et al.*[59] TDCC is a measure of correlation between parameter ranks found in two consecutive sampling experiments, which is designed to be more sensitive to agreement on the top rankings.[60] We calculated TDCC for sample size N = [100, 1000, 2000, 5000, 9000, 9600]. Starting from

N = 2000 TDCC followed a saturation trend (see Supplementary Fig. 1). Thus we estimated 10 000 as a sufficient number of Sobol's points for our analysis.

Discussion

The performance of a signal-transduction system ultimately depends on the dynamics of its protein-protein interactions (both regulatory and structural). When considering huge signaling networks such as that found at the neuronal synapse, the combined effect of all interactions in the system can not be predicted intuitively. Instead, mathematical modeling serves as a powerful tool for quantitative and predictive understanding of dynamic system behavior. Contemporary modeling techniques are designed with high levels of complexity in mind. Among them, rule-based modeling incorporates the molecular information at the level of protein sites, defining protein interaction rules in a simple syntax. This enables rule-based models to deal with combinatorial expansion, inevitably concomitant of a large signaling network.

The PSD is a good example of a large signaling complex, comprised of many distinct signaling cascades, operating in well-orchestrated manner. The elements of these signaling pathways, as well as membrane receptors, cytoskeleton components and enzymes are held together by the highly organized lattice structure built upon scaffold proteins. Interactions within this lattice environment are mostly underpinned by interactions between complementary domains, highly enriched in many PSD proteins.

Using the Kappa language, we developed a proof of concept rule-based model of a minimal consensus PSD. Molecular diversity was markedly reduced in comparison with real 2000–3000 molecular types described for PSDs. The model includes a minimal set of functional categories of proteins that are known to comprise the PSD. Among these are representatives of the main scaffolds, receptors, cytoskeleton components and signaling elements. All the elements of the model are considered from the point of view of agglomeration and complex composition, rather than signal propagation. However, the signaling mechanics can be superimposed over the existing model structure in subsequent revisions, as it includes potential outputs such as actin remodeling or receptor clustering in response to specific perturbation. The current model includes the main classes of kinases, reported for the synapse. It has the small GTPase and their regulatory partners, which allows model extension with MAPK cascade downstream. Application of post-translational modification to the network will enable temporal modeling of signal processing and allow tracking the step by step dynamics of complex formation.

We simulated the model with a stochastic algorithm, looking at the distribution of the protein clusters/snapshots at the time point when the system was reaching a steady state. Where available we used the stoichiometry data, derived from literature to set the initial conditions. As the exact kinetic information for the selected system is very sparse and incomplete we decided to define the parameter space setting constraints to the equilibrium constants and rates of the monomolecular reactions. Further uniform sampling from the parameter hypercube allowed us to select parameter sets that favor either larger protein aggregation or those that give the small clusters. Although the postsynaptic density could be isolated as a single whole, several immunoprecipitation-based pull-down studies each revealed the different subsets of PSD proteins, with just a partial overlap between them.[1] The latter might indicate that some bonds within the complex are weaker and some are stronger, so that the real system may exist as a combination of subclusters/'signalosomes'. From this point of view the parameter sets giving the collection of the moderate-sized complexes may actually be realistic also.

The current model does not include the cooperativity in domain–domain association; all domains are considered to bind unconditionally and independently. This is not what we observe in the real system. Indeed, the domains can influence each other by steric blocking or changing the protein conformation. However we found this simplification reasonable when looking at the capacity of the system components to aggregate and the properties of these aggregations. We acknowledge that future work will inevitably require adding the regulatory phenomena to the domain–domain interactions to make the model more biologically plausible.

Even the minimal model of the PSD presented here still has an impressively large number of parameters ($\sim$90). This makes even this minimal system barely computationally tractable thus, further simplification is still appealing. We assigned single constants at the level of domain–domain interactions rather than individually by protein pair. This significantly reduced the parameter space but of course we then have to rely on sensitivity analysis to reveal network critical interactions that would benefit from more accurate parameters. To achieve this we ranked the parameters according to the relative impact on cluster size and diversity. The analysis allows us to identify those constants that have most influence on the complex size and composition. Conversely, parameters found as 'important' need more extensive investigation with potential splitting into more specific/important ones. Our analysis identified the set of most important parameters, largely corresponding to interactions between the different classes of scaffold proteins: GK- GK binding domain, SH3-Proline-rich domain, *etc*. Although these results might seem quite predictable we also found k37, the association constant for coiled-coil domain as the very influential for both size and composition of the complexes that needs the further investigation.

We simulated a knockout phenotype of a PSD protein, changing the concentration of its main scaffold protein PSD95 to 0. We observed large differences in the sizes of complexes produced by the simulations and this corresponds to the main structural role of PSD95. The remaining three MAGUK proteins were still available in simulation, but due to stoichiometric ratio (300 PSD95 : 50 PSD93 : 37 SAP 102 : SAP97)[22] their impact is not as strong as of PSD95. All the biggest complexes of 'wild-type' simulation composed of long polymers of self-associated PSD95 and MAGUK proteins, each bound to some interacting partners. The biggest complexes obtained in 'mutant' simulation comprised of the polymers of SHANK, which hold the rest of the system components. As the Shank appears twice less abundant than PSD95 (150 copies/average PSD), but still in good excess compared to other scaffold proteins it is a candidate substitute for the structural role of PSD-95. The model is not limited to gross knockout mutant

phenotypes but can also simulate many other more subtle perturbations. If the different splice variants of the same protein have different binding domains, this information could be easily introduced to the model. In similar way, the specific drug application could be introduced to the model if it alters the certain domain–domain interactions.

It was interesting to directly compare the topology of the static PPI network with that obtained through the simulations. The static interaction model, albeit with a reduced/minimal protein complement shares the same general network topology and features as that observed in published receptor models.[2] Our simulated rules-based models graphs appeared to have more of a loose tree-like structure. The small world feature of the condenses static interaction networks appears, perhaps not unsurprisingly to be more complex in nature with a distribution of shortest paths for each pair of molecules within the overall network architecture. Further, no single simulated network even contained all 54 proteins (mostly in the range of 20–30 proteins). The topological structure we obtain from the simulated approach is clearly dependant upon, and varies with the interaction affinities and stoichiometry for the same given set of proteins. At the same time, model presented here is clearly lacking in terms of the molecular classes available for constructing the network (54 proteins compared with hundreds-thousands at real synapses). This also might explain the sparse connectivity observed when we optimize the simulation conditions towards larger complexes. Having a protein-protein interaction map and protein domain structure one would be able to generate the bigger models on the fly, just choosing the appropriate rule from the list. Adding the post-translational modifications will enable us to build a more dynamic picture with further analysis of different parameter influence on signal propagation.

Conclusions

The work presented here illustrates the possibility of extending a qualitative protein-protein interaction map into a quantitative executable model. Existing static PPi models of synapse proteome cannot capture the dynamic complexity and subtle perturbations of molecular structure we expect to find in the postsynaptic density. We developed the rule-based model that predicts the quantitative distribution of protein complexes with realistic structure and composition, thus, having a big potential in providing molecular mechanisms of physiological phenomena at the post synaptic level.

This work has made use of the resources provided by the Edinburgh Compute and Data Facility (ECDF) (http://www.ecdf.ed.ac.uk/). The ECDF is partially supported by the eDIKT initiative (http://www.edikt.org.uk). The research leading to these results has received funding from the European Union Seventh Framework Programme under grant agreement nos. HEALTH-F2-2009-241498 (“EUROSPIN” project) and HEALTH-F2-2009-242167 (“SynSys-project”).

Notes and references

1 M. O. Collins, H. Husi, L. Yu, J. M. Brandon, C. N. G. Anderson, W. P. Blackstock, J. S. Choudhary and S. G. N. Grant, *J. Neurochem.*, 2006, **97**, 16–23.
2 A. J. Pocklington, M. Cumiskey, J. D. Armstrong and S. G. N. Grant, *Mol. Syst. Biol.*, 2006, **2**, 2006.0023.
3 E. Fernández, M. O. Collins, R. T. Uren, M. V. Kopanitsa, N. H. Komiyama, M. D. R. Croning, L. Zografos, J. D. Armstrong, J. S. Choudhary and S. G. N. Grant, *Mol. Syst. Biol.*, 2009, **5**.
4 A. Zeke, M. Lukács, W. A. Lim and A. Remény, *Trends Cell Biol.*, 2009, **19**, 364–374.
5 W. R. Burack and A. S. Shaw, *Curr. Opin. Cell Biol.*, 2000, **12**, 211–216.
6 E. Kim and M. Sheng, *Nat. Rev. Neurosci.*, 2004, **5**, 771–781.
7 J. F. Sturgill, P. Steiner, B. L. Czervionke and B. L. Sabatini, *J. Neurosci.*, 2009, **29**, 12845–12854.
8 B. Xiao, J. C. Tu and P. F. Worley, *Curr. Opin. Neurobiol.*, 2000, **10**, 370–374.
9 C. Nourry, S. G. N. Grant and J.-P. Borg, *Sci. STKE*, 2003, **179**, re7.
10 T. Pawson and J. D. Scott, *Science*, 1997, **278**, 2075–2080.
11 W. Feng and M. Zhang, *Nat. Rev. Neurosci.*, 2009, **10**, 87–99.
12 N. L. Novère and T. S. Shimizu, *Bioinformatics*, 2001, **17**, 575–576.
13 K. M. Franks, T. M. Bartol and T. J. Sejnowski, *Neurocomputing*, 2001, **38**, 9–16.
14 S. S. Andrews, N. J. Addy, R. Brent and A. P. Arkin, *PLoS Comput. Biol.*, 2010, **6**, e1000705.
15 V. Danos, J. Feret, W. Fontana, R. Harmer and J. Krivine, *TRANSACTIONS ON COMPUTATIONAL SYSTEMS BIOLOGY XI, Lecture Notes in Computer Science*, 2009, **5750**, 116–137.
16 W. S. Hlavacek, J. R. Faeder, M. L. Blinov, R. G. Posner, M. Hucka and W. Fontana, *Science's STKE*, 2006, **344**, re6.
17 D. Bray and R. B. Bourret, *Mol. Biol. Cell*, 1995, **6**, 1367–1380.
18 K.-H. Lee, A. R. Dinner, C. Tu, G. Campi, S. Raychaudhuri, R. Varma, T. N. Sims, W. R. Burack, H. Wu, J. Wang, O. Kanagawa, M. Markiewicz, P. M. Allen, M. L. Dustin, A. K. Chakraborty and A. S. Shaw, *Science*, 2003, **302**, 1218–1222.
19 P. J. Woolf and J. J. Linderman, *J. Theor. Biol.*, 2004, **229**, 157–168.
20 V. Danos and L. J. Schumacher, *Theor. Comput. Sci.*, 2009, **410**, 1003–1012.
21 V. Danos, J. Feret, W. Fontana and J. Krivine, *Proceedings of APLAS*, 2007.
22 M. Sheng and C. C. Hoogenraad, *Annu. Rev. Biochem.*, 2007, **76**, 823–847.
23 X. Chen, L. Vinade, R. D. Leapman, J. D. Petersen, T. Nakagawa, T. M. Phillips, M. Sheng and T. S. Reese, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 11551–11556.
24 J. D. Petersen, X. Chen, L. Vinade, A. Dosemeci, J. E. Lisman and T. S. Reese, *The Journal of Neuroscience*, 2003, **23**, 11270–11278.
25 Y. Lin, V. A. Skeberdis, A. Francesconi, M. V. L. Bennett and R. S. Zukin, *J. Neurosci.*, 2004, **24**, 10138–10148.
26 M. SHENG and D. T. PAK, *Ann. N. Y. Acad. Sci.*, 1999, **868**, 483–493.
27 H. Shin, Y.-P. Hsueh, F.-C. Yang, E. Kim and M. Sheng, *The Journal of Neuroscience*, 2000, **20**, 3580–3587.
28 M. K. Baron, T. M. Boeckers, B. Vaida, S. Faham, M. Gingery, M. R. Sawaya, D. Salyer, E. D. Gundelfinger and J. U. Bowie, *Science*, 2006, **311**, 531–535.
29 K. Hirao, Y. Hata, N. Ide, M. Takeuchi, M. Irie, I. Yao, M. Deguchi, A. Toyoda, T. C. Sudhof and Y. Takai, *J. Biol. Chem.*, 1998, **273**, 21105–21110.
30 R. J. Wenthold, K. Prybylowski, S. Standley, N. Sans and R. S. Petralia, *Annu. Rev. Pharmacol.*, 2003, **43**, 335–358.
31 C. L. Waites, C. G. Specht, K. Härtel, S. Leal-Ortiz, D. Genoux, D. Li, R. C. Drisdel, O. Jeyifous, J. E. Cheyne, W. N. Green, J. M. Montgomery and C. C. Garner, *J. Neurosci.*, 2009, **29**, 4332–4345.
32 L. Chen, D. M. Chetkovich, R. S. Petralia, N. T. Sweeney, Y. Kawasaki, R. J. Wenthold, D. S. Bredt and R. A. Nicoll, *Nature*, 2000, **408**, 936–943.
33 J. C. Tu, B. Xiao, S. Naisbitt, J. P. Yuan, R. S. Petralia, P. Brakeman, A. Doan, V. K. Aakalu, P. F. Worley, A. A. Lanahan and M. Sheng, *Neuron*, 1999, **23**, 583–592.
34 G. Krapivinsky, I. Medina, L. Krapivinsky, S. Gapon and D. E. Clapham, *Neuron*, 2004, **43**, 563–574.
35 V. Pena, M. Hothorn, A. Eberth, N. Kaschau, A. Parret, L. Gremer, F. Bonneau, M. R. Ahmadian and K. Scheffzek, *EMBO Rep.*, 2008, **9**, 350–355.

36 R. D. Rogge, C. A. Karlovich and U. Banerjee, *Cell*, 1991, **64**, 39–48.
37 J. E. Sommer and E. C. Budreck, *J. Neurosci.*, 2009, **29**, 5367–5369.
38 X.-M. Ma, J. Huang, Y. Wang, B. A. Eipper and R. E. Mains, *The Journal of Neuroscience*, 2003, **23**, 10593–10603.
39 C. A. Rabiner, R. E. Mains and B. A. Eipper, *Neuroscientist*, 2005, **11**, 148.
40 M. Wyszynski, J. Lin, A. Rao, E. Nigh, A. H. Beggs, A. M. Craig and M. Sheng, *Nature*, 1997, **385**, 439–442.
41 A. Wechsler and V. I. Teichberg, *EMBO J.*, 1998, **17**, 3931–3939.
42 S. Naisbitt, E. Kim, J. C. Tu, B. Xiao, C. Sala, J. Valtschanoff, R. J. Weinberg, P. F. Worley and M. Sheng, *Neuron*, 1999, **23**, 569–582.
43 M. A. McNiven, L. Kim, E. W. Krueger, J. D. Orth, H. Cao and T. W. Wong, *J. Cell Biol.*, 2000, **151**, 187–198.
44 A. J. Robison, M. A. Bass, Y. Jiao, L. B. MacMillan, L. C. Carmody, R. K. Bartlett and R. J. Colbran, *J. Biol. Chem.*, 2005, **280**, 35329–35336.
45 Z. Xie, D. P. Srivastava, H. Photowala, L. Kai, M. E. Cahill, K. M. Woolfrey, C. Y. Shum, D. J. Surmeier and P. Penzes, *Neuron*, 2007, **56**, 640–656.
46 I. M. Brooks and S. J. Tavalin, *J. Biol. Chem.*, 2011, **286**, 6697–6706.
47 B. Y. Chang, M. Chiang and C. A. Cartwright, *J. Biol. Chem.*, 2001, **276**, 20346–20356.
48 M. M. Rodriguez, D. Ron, K. Touhara, C.-H. Chen and D. Mochly-Rosen, *Biochemistry*, 1999, **38**, 13787–13794.
49 A. Foster-Barber and J. M. Bishop, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 4673–4677.
50 N. Takagi, H. H. Cheung, N. Bissoon, L. Teves, M. C. Wallace and J. W. Gurd, *J. Cereb. Blood Flow Metab.*, 1999, **19**, 880–888.
51 G. Cuesto, L. Enriquez-Barreto, C. Caramés, M. Cantarero, X. Gasull, C. Sandi, A. Ferrús, Á. Acebes and M. Morales, *J. Neurosci.*, 2011, **31**, 2721–2733.
52 C. B. Chan and K. Ye, *The Scientific World JOURNAL*, 2010, **10**, 613–623.
53 N. Guhan and B. Lu, *Trends Neurosci.*, 2004, **27**, 645–648.
54 C. Sawallisch, K. Berhörster, A. Disanza, S. Mantoani, M. Kintscher, L. Stoenica, A. Dityatev, S. Sieber, S. Kindler, F. Morellini, M. Schweizer, T. M. Boeckers, M. Korte, G. Scita and H.-J. Kreienkamp, *J. Biol. Chem.*, 2009, **284**, 9225–9236.
55 J. E. Brenman, K. S. Christopherson, S. E. Craven, A. W. McGee and D. S. Bredt, *The Journal of Neuroscience*, 1996, **16**, 7407–7415.
56 J. W. Gurd and N. Bissoon, *J. Neurochem.*, 1997, **69**, 623–630.
57 D. Wendholt, C. Spilker, A. Schmitt, A. Dolnik, K.-H. Smalla, C. Proepper, J. Bockmann, K. Sobue, E. D. Gundelfinger, M. R. Kreutz and T. M. Boeckers, *J. Biol. Chem.*, 2006, **281**, 13805–13816.
58 H. Maruoka, D. Konno, K. Hori and K. Sobue, *J. Neurosci.*, 2005, **25**, 1421–1430.
59 S. Marino, I. B. Hogue, C. J. Ray and D. E. Kirschner, *J.Theor. Biol.*, 2008, **254**, 178–196.
60 R. L. Iman and W. J. Conover, *Technometrics*, 1987, **29**, 351–357.
61 G. Csárdi and T. Nepusz, *Inter.Journal Complex Systems*, 2006, **1695**, 2006.
62 G. Pujol and B. Iooss, 2008.
63 V. Danos, J. Feret, W. Fontana and J. Krivine, *Lect. Notes Comput. Sci.*, 2008, **4905**, 83–97.
64 C. Dutang, 2009.
65 D.T. Lang, 3.1-0 edn, 2010, pp. This package provides many approaches for both reading and creating XML (and HTML) documents (including DTDs), both local and accessible *via* HTTP or FTP. It also offers access to an XPath ''interpreter''.